

# Built to Last: Twenty-Five Years of Delivering Digital Infrastructure

Twenty-five years of building digital infrastructure that governments and global institutions rely on. Grounded in the open data movement, responding to the AI challenge.

A Link Digital Whitepaper published to mark our 25th year, November 2026

# Table of contents

1. Introduction
2. A milestone and a shift
3. Built to last: The lessons of twenty-five years
4. Creating digital infrastructure: The three jobs of a living platform
5. The Link Digital approach: Stewards, not just builders
6. The next twenty-five years: The next twenty-five years: Making data trustworthy in the machine-readable era
7. Work with us
8. About Link Digital

# 1. Introduction

As Link Digital celebrates twenty-five years as a company, a significant shift is taking place in the world of digital infrastructure: the move to public code. This is software source code, algorithms, and civic policies designed by and for public sector organisations published under open license and designed to serve as digital public infrastructure. Open source is playing a key role in this. In addition to providing better value for money than proprietary software, the ability to inspect code, verify provenance, and avoid vendor lock-in has turned open source from a possible alternative into a default consideration, especially for the public and government technology sectors.

But adopting open source is only the beginning, not the end. A data platform, whether a public open-data portal or an internal data hub, is not a project that finishes at launch. It is a living system. Left alone, its capabilities degrade: it drifts out of support, falls behind on security updates, and it becomes vulnerable to threats that did not exist when it was built.

```

    ( value )
    25 : 500 .2 300 )
    (  => ;
}

```

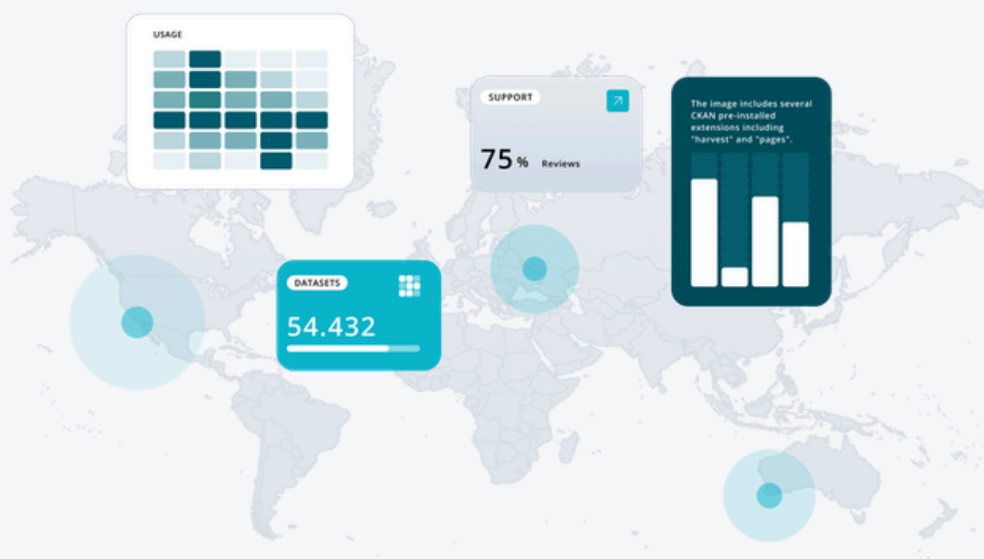
Link Digital sits at the intersection of open source, public code, and digital public goods. We are an official co-steward and global expert in the [Comprehensive Knowledge Archive Network \(CKAN\)](#), and an expert in complementary open source software (OSS) such as Drupal, as well as in the support, hosting, maintenance, migration, customisation, and upgrades of digital infrastructure, including open and internal data portals.

For a quarter of a century, our core mission has been supporting organisations to deploy the policy and broader public good gains from the data they have and create meaningful change. We are a small company but we have sought to have a much bigger impact, positioning ourselves to add value to the new ways that open data is being utilised: in innovative methods of working for government and public servants; in building trust in public institutions and combating malicious large-scale disinformation campaigns; in fighting for the continued importance of open government, including through our participation in global organisations such as the [Open Government Partnership](#).

And now we are adapting and growing our services to make open data more AI ready, providing better quality data to help build models of AI that are more trustworthy, safer, diverse and socially constructive.

In many respects, an anniversary is just a circled number on a calendar. What is more important is Link Digital's commitment to staying the course by continuing to support openness and its civic and human benefits and refusing to abandon this mission in favour of potentially easier commercial pivots. We want to continue to serve as a necessary and consistent force in an era where social and civic discourse is increasingly fractured, and the experience of the Internet has diminished as user interfaces and web experiences become homogenised and siloed into dominant tech platforms.

We understand where the open source and digital public infrastructure world is heading. There is a much that can be said about this but, in the current conditions, our work can essentially be boiled down to three key tasks:



## Protect it



Automated traffic now accounts for the majority of web traffic, and AI-driven traffic is growing far faster than human traffic. Public data portals are directly exposed to this load.

## Enhance it



You will have heard the term 'agentic' being used a lot in relation to how data portals are evolving. An agentic data portal replaces passive repositories with autonomous AI agents. This allows it to perform functions such as interpreting business goals, extracting raw data across disparate systems, automatically cleaning and joining schemas, and executing complex analytics, delivering actionable insights on demand. This is the direction we are moving in with our new Ask AI feature. While Ask AI does not make a traditional portal fully agentic, it transforms it into an agent-ready, conversational semantic search layer.

## Sustain it



Open source platforms release updates regularly. Staying current on major versions is what keeps a platform secure, supported, and performant.

The following paper explains the discipline of Link Digital's work, illustrating it with verified projects we have done for clients. It will also look ahead to where trust in data is heading next and at future activities we have planned.

## 2. A milestone and a shift

Twenty-five years is a very long time in technology. Platforms that looked permanent have disappeared and standards have risen and fallen.

When Link Digital was first established as Link Web Services in 2001, we closed an essential gap between technical and creative services that were often delivered by different teams. We worked to link the various web related services to deliver websites, interactive PDFs, multimedia kiosks, CD-ROMS and multimedia digital communications.

In 2004, we updated our branding to Link Digital as a way of better grounding ourselves as a digital service agency of choice in the local and interstate market in Australia. Amazon Web Services established a presence in Australia in 2012 and Link Digital was one of its first local Advanced Tier Consulting Partners. The same year, Link Digital designed, built and hosted one of the Australian national government's early open data portals, [data.gov.au](https://data.gov.au).

We now build data management solutions for clients globally, using CKAN and other open source tools, and are positioned at the operational intersection of public code, OSS and digital public infrastructure, helping to turn open-source principles into sustainable public infrastructure.

And while our core mission remains empowering our clients to make meaningful change with their data, we are using our extensive experience to think differently about our work and how to make open data infrastructure more effective in the face of advances in machine learning. This includes creating agentic portals - data management infrastructure that incorporates AI agents to understand, manipulate and act upon data on a user's behalf. In a traditional portal, users rely on key words to find a dataset. By using semantic understanding portals don't just use filenames; they understand the context of the data. We now incorporate natural language or semantic search interaction into our open data portals and catalogues.

This work seeks to rectify a key tension at the heart of contemporary data governance: many organisations now sit on such large collections of data that it has become functionally inaccessible to use by non-technical or expert staff. It is also part of a larger shift that Link Digital has been an integral part of, described by some observers as the '[Fourth Wave of Open Data](#)'. [The Third Wave of Open Data](#) involved the proposition that the open data ecosystem had moved beyond focusing on simplifying data and its impacts, to how governments, private corporations and civil society organisations can better know and understand the data they have and share it to provide improved products and services, and a more informed citizenry. In the Fourth Wave, the focus is on making open data more conversational and AI ready.



We'll talk more about what Link Digital is doing in this space later later in this document. For now, it is important to stress that while there have been enormous changes during Link Digital's 25 years of existence, we believe three truths hold.

## 01

---

The first truth is that organisations need to publish, manage, and share data, and they need to keep doing it reliably for years, not months.

## 02

---

The second truth is that although some of the tools and approaches Link Digital uses to help organisations do this may be evolving, we continue to rely on OSS as our main building tool and to adhere to open source principles more generally.

OSS provides flexibility, scalability and cost effectiveness. OSS is not cost-free, but it can be utilised at a significantly lower price than proprietary software options. And, unlike proprietary systems that cannot be inspected, open source platforms let organisations audit code, verify where their data lives, and avoid dependence on a single vendor.

Open source technology will play a major role in the European Union's recently announced European Technological Sovereignty Package, a set of measures to strengthen Europe's technology sovereignty by increasing capacity in areas like semiconductor production, AI, and cloud computing. And OSS has been recognised as a core part of critical infrastructure security in United States cybersecurity planning.

There are also numerous examples of, particularly, government's shifting their operations from proprietary to OSS.

The German state of Schleswig-Holstein recently [initiated a landmark shift in its public sector IT strategy by transitioning roughly 30,000 government workstations off proprietary solutions in favour of OSS](#). The primary drivers behind this move are regaining digital sovereignty, data privacy and compliance, and long-term cost efficiencies.

And earlier this year, the French Government announced [plans to migrate 2.5 million civil servant workstations away from proprietary tech to OSS](#). The move, which encompasses collaboration tools, antivirus software, AI platforms, databases, virtualisation systems, and network equipment, was made on the grounds of security and data sovereignty concerns.

## 03

---

The third truth - and this is the point the migration conversation usually misses - is that moving to open source is a decision about ownership, not a finish line. Once an organisation owns its platform, it also owns the responsibility to keep it running.

A data platform is a living system. This rule applies whether the platform is a public open-data portal serving citizens or an internal data hub serving staff, or a combination of the two. Both need protecting, enhancing, and sustaining. Both reward the organisation that treats the platform as something to be kept alive and penalise the one that treats it as something that was finished.



## 3. Built to last: The lessons of twenty-five years

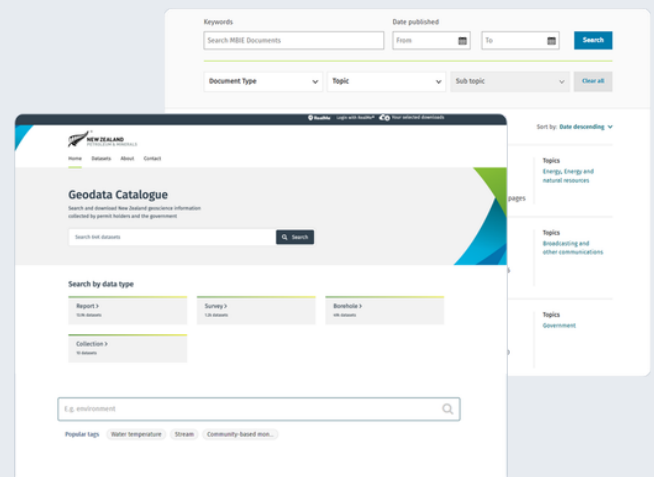
Executed well, a migration – whether it involves the switch from proprietary or another open source software – can have a significant beneficial impact for an organisation.

Let's look at some examples from the digital infrastructure work that Link Digital has done for clients globally.



# New Zealand Ministry of Business, Innovation and Employment

The [Resources Markets Branch](#) of New Zealand's [Ministry of Business, Innovation and Employment \(MBIE\)](#) is the country's regulatory agency for petroleum and minerals resources.



It has responsibility for collecting and managing petabytes of geoscience data, including that gathered under mandatory industry reporting, including spatial information, working with our treaty partners, the tangata whenua of Aotearoa New Zealand. MBIE's previous online data storage and management systems – operated using proprietary software – were not meeting the rapidly changing needs of users and stakeholders. MBIE was data rich but found it difficult to make connections and gain insights from data in its possession.

Link Digital helped MBIE replace this setup with two interconnected CKAN implementations: an internal asset register for internal data management and [a public-facing Geodata Catalogue](#). These were connected to each other via the CKAN API and integrated with multiple systems involved in MBIE's data pipeline and workflow processes: Azure cloud infrastructure, RealMe (New Zealand's digital ID platform), Azure Active Directory SSO, and geospatial tools like ArcGIS and FME Workbench. The project also involved significant data migration challenges, moving data from legacy systems to the new CKAN platform.

Azure cloud infrastructure

Azure Active Directory SSO

RealMe (New Zealand's digital ID platform)

ArcGIS

FME Workbench

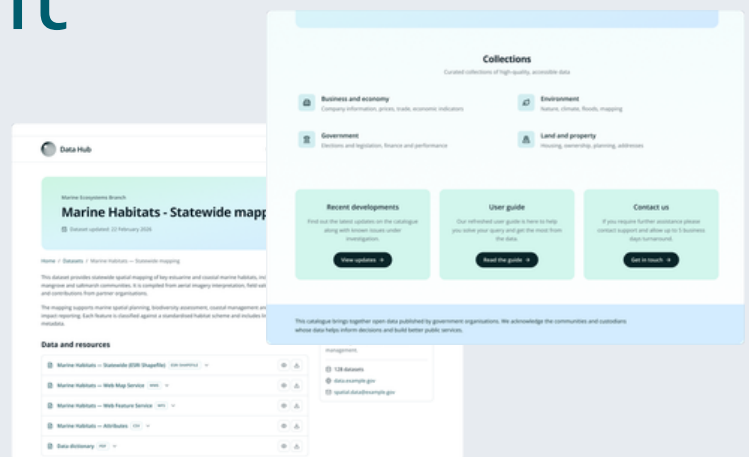
The new system has significantly improved data discoverability, with previously hard-to-find data now easily accessible to users.

**Learn more about our work on this project here:**

**Read more**

# Upgrading a national OECD government open data portal to a full CKAN environment

The portal, the Government's central open catalogue for the discovery of public data, has been operating for over a decade.



As the national data ecosystem evolved and the portal transitioned from one statutory authority to another, Link Digital was engaged to securely transition the portal exclusively to CKAN to make it simpler to administer and reduce maintenance costs, as opposed to its previous mixed software architecture in which CKAN was only one part of the tech stack.

Maintaining all the datasets native to the portal while reintegrating the harvester process seamlessly back into CKAN was challenging due to the outdated and insecure versions of the software used by the previous authority overseeing it. We worked with the new owner to fully upgrade CKAN and migrate the site to a new domain, while ensuring total data integrity and continuity. While the portal was being migrated there was only a brief period of a few days during which new datasets could not be uploaded into the portal, but otherwise it remained fully online and accessible to the public.

Other benefits of the new site include:

- 01** A new federal data architecture that includes a harvesting mechanism designed to identify and communicate the original source of data sets to users, effectively managing potential duplicate data from other government open data portals.
- 02** A new, dynamic front page using Drupal, designed to be highly customisable so it can meet evolving requirements.
- 03** Metadata structures that conform to open standards, like DCAT.
- 04** Adhering to strict accessibility standards.

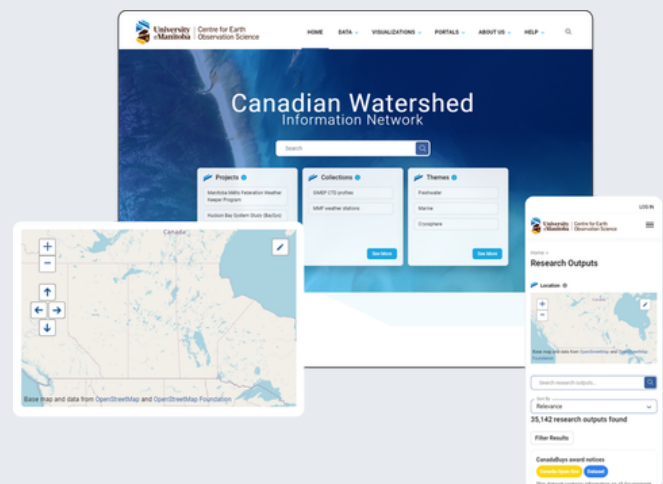
Link Digital continues to provide managed support services.

**Read more about the project here:**

[Read more](#)

# The Canadian Watershed Information Network

The Canadian Watershed Information Network (CanWIN) is a Canadian spatial research data infrastructure system hosted at the University of Manitoba; and managed by the Centre for Earth Observation Science (CEOS) within the Faculty of Environment, Earth and Resources.



A leading multidisciplinary and collaborative research centre focused on understanding how the Earth will respond to climate change, CEOS researchers conduct their fieldwork globally, with the Arctic freshwater marine system being a particular focus.

CanWIN was originally developed by the Canadian federal government to share freshwater research, using Visual Studio 2008, which ultimately was not able to capture and share the complexity of the multi-disciplinary data produced by CEOS's researchers. The CEOS realised the CanWIN ecosystem required a non-structured database that was flexible, modular, had an easy-to-use user interface, and was open-source.

CEOS was also keen for the portal redesign to make all CanWIN's data more easily discoverable by users in a user-intuitive and easy way. In 2020, CEOS engaged [Link Digital](#) to completely overhaul the portal using CKAN as its backbone, along with Drupal for content management.

The new CanWIN site has improved the ability of researchers and scientists across institutions to collaborate on climate data analysis and share findings for analysis from a hyperlocal to global scale. The new site has also delivered:



Customised metadata templates to improve data discoverability.



Greater data storage capability.



New visualisation tools to enable researchers to tell more complex stories about the data in their work.



Added user permissions to provide data sovereignty to Indigenous groups to give them granular control over their data.

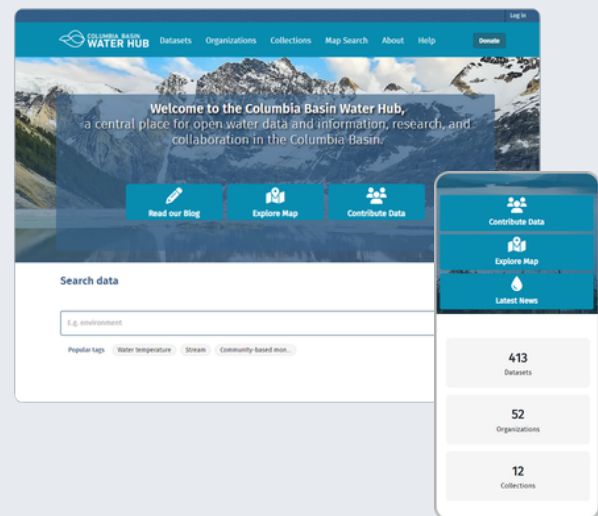
Link Digital provides the CanWIN site with ongoing support and maintenance.

**Read more about Link Digital's work on the Canadian Watershed Information Network.**

[Read more](#)

# Living Lakes Canada

Living Lakes Canada (LLC) is a non-government organisation with 20 plus years' experience working with community groups to protect freshwater within the Columbia Basin, North America's fourth largest freshwater basin.



The [Columbia Water Basin Hub](#) is LLC's centralised on-line open repository for water data collected from across the Basin. The Hub was built in partnership with [Link Digital](#), using CKAN as its software backbone, and launched in early 2021. It relies on a large network of groups that undertake community-based water monitoring, the results of which are uploaded – often directly by the groups themselves – into the Hub. This includes commercial interests, Indigenous groups, local conservation organisations, and river users such as fishers, hunters, and kayakers.

While community water monitoring was conducted in the Basin for many years prior to the Hub's establishment, there was a lack of streamlined archival and retrieval technologies for the data they collected. Information was being held on private devices and, in some cases, even on paper and was at risk of being lost.

The Hub is designed not only to store past, current and future water data, but to ensure adequate quality control and make the data more accessible and shareable.

The Hub has facilitated greatly improved accessibility to what was once inaccessible data and has become instrumental in improved decision-making processes while showcasing the tangible impact of the greater collaboration between organisations involved in water monitoring.

Other benefits include:

- 01** Interoperability with CKAN platforms run by the British Columbia and Canadian national governments.
- 02** An intuitive, user-friendly set up, making the portal relatively easy to use for local groups with varying levels of technical expertise.
- 03** The ability to create a standardised vocabulary, essential given that the data collected is diverse and collected in varied formats.
- 04** Advanced data visualisation tools.

Link Digital provides the CanWIN site with ongoing support and maintenance.

**Read more about Link Digital's work on the Canadian Watershed Information Network.**

**Read more**

# 4. Creating Digital Infrastructure: the three jobs of a living platform

If a data platform is a living system, keeping it healthy is ongoing work. In the current situation, this work focuses down into three jobs.

## **Job one:** protect it

Running a CKAN portal comes with responsibilities for the data you host and for the users and the trust they place in your platform.



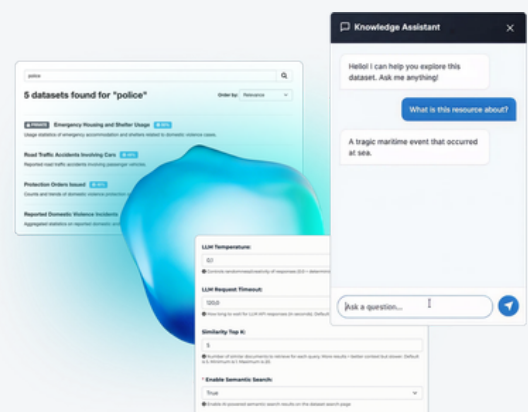
But while your focus is on publishing datasets, improving metadata, or enabling collaboration, the day-to-day demands of infrastructure management often slip under the radar: security patches, performance tuning, daily backups, and uptime monitoring. But these aren't optional. They're mission critical.

Behind every high-performing CKAN instance is often something less visible but equally important: reliable, secure, and well-managed hosting. A managed hosting partner, such as Link Digital, handles these behind-the-scenes issues, so your team can stay focused on user needs, data integrity, and insights.

Another pressing problem comes from the fact that in 2025 automated traffic counted for more than half of all web traffic, up from the year before, with a large share of it malicious. For a public data portal, this is not an abstract concept. A portal exists to be crawled, queried, and downloaded, which makes it a natural target for automated load. But, left unmanaged, that load degrades performance for real users, corrupts analytics, and consumes crawl budget and server resources. Protection - whether it is rate limiting, bot detection, and traffic management at the hosting layer, all tuned to distinguish legitimate crawlers and AI agents from abusive ones - is not optional maintenance. It is a condition of keeping the portal usable. Practical protection. This work sits naturally within managed hosting, because it depends on visibility into server behaviour that a build-and-leave arrangement does not provide.

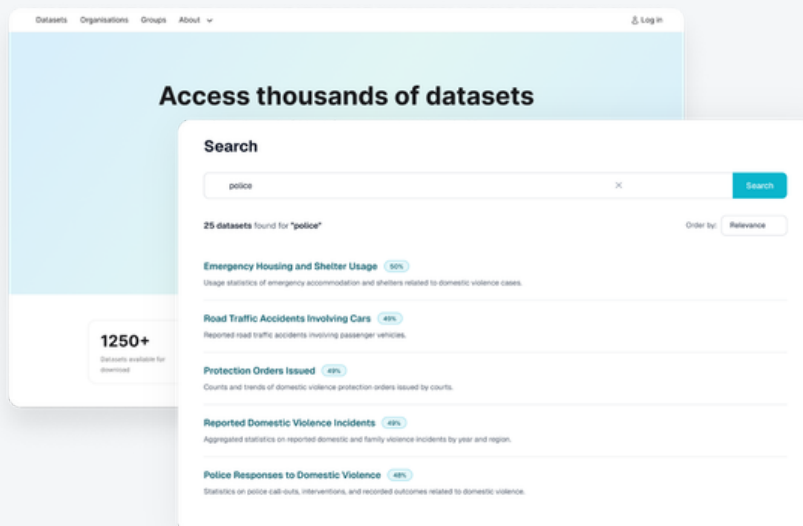
## Job two: enhance it

The reality is that publishing data is no longer enough on its own. Users increasingly expect portals to have an agentic function, i.e., to be able to ask a plain-language question and receive a reliable answer, drawn from the data the organisation already holds.



This is why Link Digital introduced its [Ask AI capability](#), which provides a grounded and secure way for users to ask questions and get verified answers from an organisation's own data - answers that are not only accurate but also accountable and traceable.

A traditional CKAN search - matching keywords in titles, descriptions, and other metadata fields - works when users know the exact phrasing used by publishers. But, for the most part, searching for data in the real world doesn't work like this. This is why Ask AI introduces semantic similarity search, which matches words, not meaning. Put simply, when a user performs a search, the system retrieves results based on semantic proximity rather than simple text overlap. This improves dataset discoverability without requiring data portal publishers and administrators to change their workflows. And to ensure accuracy, every AI-generated answer provides a direct link to the source dataset.



This not only makes the data on CKAN-powered portals more findable and accessible, but users do not need to be a data scientist to understand their search results.

Vitally, Ask AI enables users to interact with a data catalogue or portal more like they're interacting with ChatGPT, it does so without sacrificing security.



This is because Ask AI can be configured in such a way that sensitive data remains within a secure infrastructure and proprietary information is never used to train public models like ChatGPT or Gemini. It enables a data manager to define exactly which datasets the AI can access and who is authorised to query them. This means the AI cannot guess or make up facts, as it is tethered to the 'trusted data layer' in your portal.

This is enhancement in the truest sense. It does not replace the platform. It makes the data in your platform more useful to the people who need it and lowers the barrier to its effective use by allowing users to get expert level insight without needing to be a data scientist.

## Job three: sustain it

Open source platforms evolve. CKAN has a regular update schedule. For example, the most recent release, CKAN 2.12, delivers a raft of improvements to file management, DataStore performance, and the user interface.



Drupal follows its own major-version cadence. While Drupal 11 is currently the latest stable, production-ready version of the software, the next release, Drupal 12, is scheduled for December 2026, the same month that Drupal 10 reaches its end of life.



Staying current is what keeps a platform secure, supported, and fast. And the cost of falling behind on this cycle compounds. The further a platform drifts from a supported version, the harder and riskier the eventual upgrade becomes.

Sustaining a platform means treating upgrades as routine stewardship, not as emergency projects. This is why the best moment to upgrade is before an organisation is forced to, while the jump is small and planned rather than large and urgent.

## 5. The Link Digital approach: Stewards, not just builders

Plenty of firms can build a data platform. Fewer stay around for the long haul of helping the owner sustain it.

Link Digital builds data management infrastructure and builds it well. But the value we have delivered across twenty-five years comes equally from what happens after launch – the patching, protecting, enhancing, and upgrading, year after year, all delivered by an official CKAN co-steward with direct involvement in the health of the platform itself. That is the difference between a supplier and a steward. A supplier delivers and departs. A steward remains accountable for the platform staying alive.

This is why Link Digital's services are framed as an ongoing relationship rather than a one-time transaction. Support, hosting, maintenance, migration, customisation, and upgrades are not separate products. They are the continuous work of keeping a platform healthy, available to organisations that need either CKAN, or Drupal, or the two working together.



## 00. Lead & Sign

- Publish thought leadership that names the gap before the market asks
- Sign our own zero - company hours staked into working capability
- Prove it live, then ground the proof into the offer you meet at step 01



## 01. Understand & Evaluate

- Your business, requirements and needs
- Our services, processes and approach
- The unique qualities of your business



## 02. Define & Agree

- Define and agree on the project approach
- Provide you with a solution proposal
- Form an agreement that creates success for both parties



## 03. Deliver

- Onboard your team into our production processes
- Onboard our team into your business
- Form a roadmap with key stakeholders
- Execute using Agile SCRUM



## 04. Support & Enhance

- Execute solution iterations against the roadmap
- Provide clarity on feature delivery timelines through a release train
- Provide ongoing support and maintenance to the project



## 05. Evolve

- Continue enhancing beyond the roadmap to further improve
- Remain current to always provide the best possible solutions

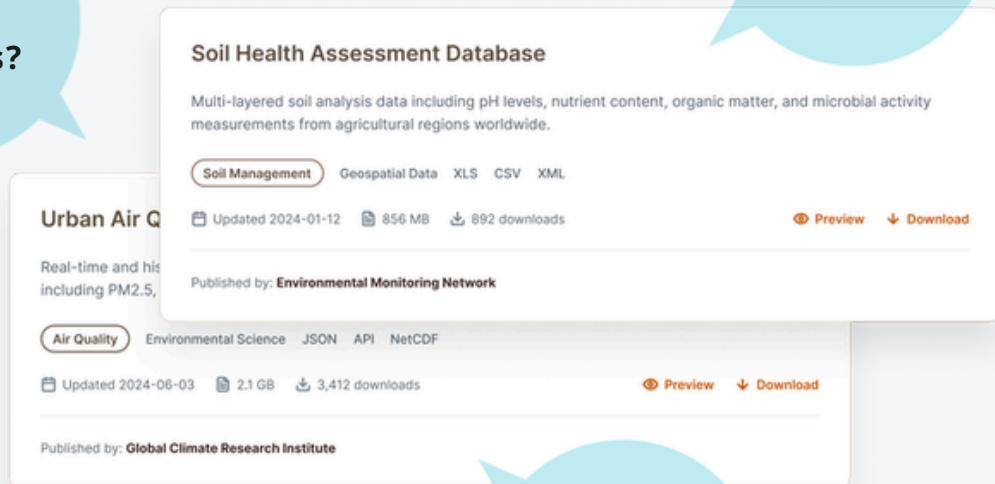
## 6. The next twenty-five years: Making data trustworthy in the machine-readable era

If the last twenty-five years were about helping organisations own their web and data platforms, a significant part of the next twenty-five will see us investing in the question of verifiable honesty and how organisations can establish and demonstrate verifiable truth and trust in their data as it moves into the AI era.

One way that we are seeking to do this is through a new framework we are building called the Objective Observer Initiative (OOI). While tools like [Croissant](#) – an open, standardised metadata format for machine learning datasets that we have been involved in developing – and Ask AI can make vital contributions to machine readability, the OOI looks at how digital systems can make information, intent and interactions objectively verifiable.

What the  
data  
contains?

Where the  
data came  
from?



How it  
changed?

A dataset can tell a machine what the data contains. Provenance can tell us where the data came from and how it changed. But those things do not necessarily tell us what the people producing the data intended to achieve, what rules governed its creation, or whether subsequent actions remained consistent with that purpose. The OOI is experimenting with what happens when that missing layer is explicitly declared and made observable.

This becomes particularly important as data is increasingly consumed by automated systems and AI. Take a community group holding a mix of reports and limited flow data on a local creek, gathered to understand why fish numbers are declining. The dataset shows what was recorded. It does not show what decision the group wants it to support, or whether future updates stay true to that purpose.

In the past, open data simply meant the file was available for download. The OOI argues that in an AI world, making data open, while still vital, is not enough because AI can't easily distinguish between high-quality facts and hallucinated or biased data, and has no reliable way to verify who created the dataset, what it was created for, or whether it's still true to that purpose.

Without a way to verify where data came from and what it was created for, AI is being built on an unstable foundation. The OOI's conjecture is that when an AI system states something is false with total confidence, part of the cause sits upstream of the model. The data it draws on rarely carries any record of how it came to exist: what someone intended when they created it, what they checked before publishing it, or whether its later use stayed consistent with that intent. Without that record, a model has no way to tell a well-supported claim from one that simply reads convincingly. The OOI's position is that data publishers have a role to play here too, by building a record of intent and verification into the data itself, before it ever reaches a model.

The screenshot displays the Link Digital website interface. The main header features 'The Provenance Sandbox' with a search bar and statistics: 495 datasets, 3 organizations, and 0 groups. Below this, there are sections for 'Organizations' (Link Digital, MLCommons) and 'OOI Anchor Family Index'. A modal window titled 'Additional Information' is open, showing metadata for a dataset.

| Field                 | Value                            |
|-----------------------|----------------------------------|
| Data last updated     | August 10, 2026                  |
| Metadata last updated | August 11, 2026                  |
| Created               | July 29, 2026                    |
| Format                | CSV                              |
| License               | Creative Commons Attribution     |
| ID                    |                                  |
| Hash                  | 2bd48adca3afd34351d12fc741926c23 |
| Same as               |                                  |
| Subresources          | Show more                        |

In addition to sharing data, Link Digital believes what is needed is verifiable honesty. The OOI is an open experiment that Link Digital is contributing to, testing how provenance and verifiable data quality can be established in a way that machines, not just people, can rely on. In a CKAN context, this means adding metadata that doesn't just describe the columns but declares the intent of the data creator. This gives AI systems a way to verify whether that data is honest, unbiased, and fit for purpose. It can also check it to see if a dataset's observed outcomes match its stated intent of the individual or organisation that uploaded the dataset.

This approach differs markedly from the current set up: feeding data into opaque machine learning tools and passively receiving the results that emerge on the other side, and where the main measurements of satisfaction and signals for consent are clicks and efficiency.

This is not a departure from what we have been doing for the last twenty-five years. It is a continuation of the same principle. Trustworthy data matters for the same reason: as decisions increasingly run through AI, the organisations that can show where their data came from and vouch for it will hold the advantage.

The OOI is already influencing Link Digital's work with data portals. As Link Digital's Executive Director Steven De Costa put it in a recent interview:



"There is an opportunity to co-construct datasets that provide optimistic signals. Such datasets would naturally be made discoverable on portals such as those we develop and maintain for our clients using CKAN, Drupal and other open technologies."

**Steven De Costa**  
Executive Director, Link Digital

# Work with us

If your organisation runs a data platform, public or internal, the questions worth asking are simple:



Is it protected against the automated traffic that now dominates the web?



Is the data in it as useful to your people as it could be?



Is it current, supported, and safe to upgrade?

Link Digital can help with any one of these, or all three, on CKAN, Drupal, or both.

## Let's talk about your data portal.

Let's talk

Email: [info@linkdigital.com.au](mailto:info@linkdigital.com.au)

Website: [linkdigital.com.au](http://linkdigital.com.au)

Canada +1 (343) 883 3051

Australia +61 (2) 6111 2907



# About Link Digital

Link Digital is a global digital consultancy and an official co-steward of CKAN, the open source data management platform. For twenty-five years, Link Digital has helped governments, research institutions, and non-profit organisations publish, manage, and share data on open source infrastructure they own and control.

## Authors and contributors

- Andrew Nette, Technical Writer
- Steven De Costa, Executive Director
- Kate Sampher, Chief Marketing Officer
- Anastasiia Shevtsova, UI/UX Design Lead
- Elli Pacis, Community Engagement Manager
- Suk Khamb-Tanganco, Head of Business Services

